

**Technical Documentation for the
2003 South Carolina Readiness Assessment of
Kindergarten and Grade-One Students**



**South Carolina Department of Education
Inez M. Tenenbaum, State Superintendent**

**Division of Curriculum Services and Assessment
Dr. Sandra Lindsay, Deputy Superintendent**

**Office of Assessment
Dr. Theresa Siskind, Director
Dr. Jim Casteel, Education Associate
Dr. Imelda C. Go, Education Associate
Joseph C. Saunders, Education Associate**

CONTENTS

LIST OF TABLES	iii
INTRODUCTION	1
CHAPTER 1. THE SCRA IN 2002–03	3
1.1 SURVEY OF TEACHERS ON THE SCRA	3
1.2 ADVISORY COMMITTEE DELIBERATIONS.....	4
Composition of the Committee	4
November 2001 Meeting	4
January 2002 Meeting.....	4
March 2002 Meeting.....	5
April 2002 Meeting.....	5
1.3 REVISIONS FOR 2002–03	5
Developmental Guidelines.....	5
Checklists	6
Reports	7
CHAPTER 2. TEST ADMINISTRATION AND SCORING	8
2.1 OVERVIEW	8
2.2 ADMINISTRATOR TRAINING	8
2.3 SOUTH CAROLINA READINESS ASSESSMENT PROFILE INTERACTIVE.....	9
History	9
Online System Features	10
Transferring Students.....	10
2.4 SCORING	11
2.5 SCORE REPORTING	11
Student Home Report.....	12
Class Roster	12
School, District, and State Summaries.....	12
CHAPTER 3. TECHNICAL CHARACTERISTICS OF THE INDICATORS AND RATINGS	14
3.1 YEAR-END PERFORMANCE INDICATORS	14
3.2 YEAR-END PERFORMANCE RATINGS	15
CHAPTER 4. VALIDITY	17
4.1 OVERALL VIEW OF DIF ANALYSIS	17
General Description	17
Standardized Mean Difference Procedure	18
NAEP DIF Classification.....	18
4.2 RESULTS OF DIF ANALYSES.....	18
APPENDIX A.....	20
APPENDIX B	23
WORKS CITED	30

TABLES

1.1	2002–03 SCRA Functional Components for Kindergarten and Grade One	6
2.1	Frequency of Performance Ratings Needed to Achieve Various Component Ratings	13
3.1	Summary of Major Statistics for Year-End Performance Indicators.....	14
3.2	Summary of Major Statistics for Year-End Performance Ratings.....	15
3.3	State Summary of Year-End Performance Ratings of the Kindergarten Assessment for First-Grade Readiness	15
3.4	State Summary of Year-End Performance Ratings of the First-Grade Assessment for Second-Grade Readiness	16
4.1	Summary of DIF Classification for Year-End Performance Indicators.....	19
4.2	Summary of DIF Classification for Year-End Performance Ratings	19
A.1	Item Statistics: Year-End Performance Indicators for Kindergarten and First Grade.....	21
A.2	Item Statistics: Year-End Performance Ratings for Kindergarten and First Grade.....	22
B.1	DIF Data: Year-End Performance Indicators for Kindergarten and Gender	24
B.2	DIF Data: Year-End Performance Indicators for Kindergarten and Ethnicity	25
B.3	DIF Data: Year-End Performance Indicators for First Grade and Gender	26
B.4	DIF Data: Year-End Performance Indicators for First Grade and Ethnicity	27
B.5	DIF Data: Year-End Performance Ratings for Kindergarten and Gender.....	28
B.6	DIF Data: Year-End Performance Ratings for Kindergarten and Ethnicity	28
B.7	DIF Data: Year-End Performance Ratings for First Grade and Gender	29
B.8	DIF Data: Year-End Performance Ratings for First Grade and Ethnicity	29

INTRODUCTION

The South Carolina Readiness Assessment (SCRA) is a state-level assessment for kindergarten and grade-one students. It replaced the Cognitive Skills Assessment Battery (CSAB), which had been used in the state of South Carolina since 1978. Unlike the CSAB, which was administered in the first month of the first grade and was a single, high-stakes test, the SCRA is a system of continuous assessment that documents a student's performance through focused observations of daily classroom activities. The CSAB was used for the last time in fall 2001.

The development of the SCRA began in 1999 when the South Carolina Department of Education (SDE) awarded a contract to Harcourt Educational Measurement (HEM) to develop a statewide continuous assessment system that would provide a better understanding of student readiness for school. The HEM submitted a proposal to use a version of the Work Sampling System (WSS) modified specifically for the state of South Carolina.

The WSS, used in many states, is a curriculum-embedded continuous assessment process where teachers observe their students during everyday classroom activities to gain a fuller picture of the students' development. Focusing on three key areas—English language arts, mathematics, and personal and social development—these observations can provide a clearer understanding of a child's readiness for school. The modified WSS is aligned with the South Carolina curriculum standards.

The WSS is the major product of Rebus, the company established by Samuel Meisels to develop, market, and support this assessment system. Since 1991, Rebus has been able to accomplish this objective in all fifty states, the District of Columbia, and Canada (HEM 1999, 25). The Department of Defense Education Agency (a civilian agency of the U.S. Department of Defense) adopted the program for all of the kindergarten through third-grade classrooms worldwide. During the 1990s, a number of states, including Pennsylvania and Maryland, started with a small sample of their schools using the system. After initial positive responses in the District of Columbia, for example, administrators in that system decided to implement the WSS in all 116 of their elementary schools over a six-year period (Monrad and Mandeville 1996).

Pearson bought Rebus in 2000. The name Rebus still exists and is mostly associated with the WSS. In 1999, the SDE granted the HEM the contract to develop a readiness assessment for students in kindergarten and grade one. The HEM has contracted with Rebus to use the WSS. However, the HEM is providing the scoring services, using its own in-house facilities.

The SCRA is an ongoing assessment of students in kindergarten and grade one. Its mechanism is a checklist that is completed by teachers on three occasions during the school year. Only the last checklist assessment is used for reporting students' performance. The checklist is organized on the basis of three subject domains, which are labeled "Personal and Social Development," "English Language Arts," and "Mathematics." Each of these domains is divided into a series of functional components under which are sets of performance indicators that are based on the South Carolina curriculum standards. During workshops that are normally conducted in the summer, teachers receive training in completing the checklist according to a set of developmental guidelines.

The first statewide SCRA field test took place during the 2000–01 school year. The system centered on the use of a checklist of developmental indicators that kindergarten and first-grade teachers completed three times during the school year for each of their students. The two checklists, one for students in kindergarten and one for those in grade one, were accompanied by development guidelines designed to give all teachers a similar understanding of the checklist items. Minor revisions were made to the system on the basis of this first pilot test.

The SCRA was field-tested on a larger scale during the 2001–02 year. An SCRA advisory committee was also established in November 2001. This group consisted of teachers and administrators from schools around the state, SDE administrators, and personnel from both the HEM and Rebus, Inc., the company that created the version of the WSS on which the SCRA is based. The advisory committee has continued to meet regularly and has recommended a number of substantial changes to the SCRA. The committee made major revisions for 2002–03, including changes both to the checklists and to the guidelines.

The first section of this report, chapters 1 and 2, provides an overview of the major changes made to the 2002–03 checklists and to the administration and the scoring of the SCRA. The second section, chapters 3 and 4, documents the technical characteristics of the test items. Detailed information about the evolution of readiness assessment in South Carolina, the WSS and its being modified to become the SCRA, and technical aspects of previous SCRA administrations can be found in an earlier technical report (Huynh, Prior, and Gallant-Taylor 2004).

Chapter 1

THE SCRA IN 2002–03

This chapter describes how the 2001–02 SCRA teacher survey administered to kindergarten and grade-one teachers and the 2001–02 SCRA advisory committee deliberations affected the 2002–03 SCRA administration.

1.1 SURVEY OF TEACHERS ON THE SCRA

In an effort to continue their evaluation of the SCRA in prior years, SouthEastern Regional Vision for Education (SERVE) developed and sent surveys to South Carolina kindergarten and grade-one teachers. The survey was developed as a result of teacher focus groups. At the January 2002 SCRA advisory committee meeting, copies of the drafted SCRA teacher survey were made available for committee members' review. Committee members recommended changes in the method of distributing the survey and in the wording of survey items; they also added some survey items. The survey was piloted with teachers in the committee members' respective districts. On the basis of the survey pilot, the committee was to make recommendations to SERVE.

The final version of the SCRA teacher survey consisted of forty items. Thirty-five items were multiple choice, three items were open-ended, and two items were on a continuum scale. The survey was divided into four major sections: "Completing the SCRA" (seventeen items), "Using Data from the SCRA" (six items), "Training and Support for the SCRA" (ten items), and "Respondents to the Survey" (seven items).

In February 2002, the SCRA teacher surveys were sent to districts for distribution to their kindergarten and grade-one teachers. The deadline for the completion and return of surveys was March 22, 2002. There were 4,540 surveys returned. Preliminary results were available in April for the advisory committee's meeting. These results included responses to questions such as these:

- "How often do you think the SCRA rating should be completed during the school year?"
- "What types of documentation do you think teachers should be *required* to keep to support the ratings they give children? Mark all that apply."
- "When do you think it would be most helpful to receive the reports?"
- "The SCRA has impacted my district in the following ways: mark all that apply."

1.2 ADVISORY COMMITTEE DELIBERATIONS

Composition of the Committee

The SCRA advisory committee, which was formed in November 2001, has twenty-five members. Eleven of them are school district personnel: two from the Greenville County School District and one each from Spartanburg School District Three, York School District One, the Berkeley County School District, Florence School District One, the Charleston County School District, Richland School District One, Greenwood School District Fifty, Lexington School District One, and Lexington School District Five. Three are faculty members at the University of South Carolina, Furman University, and Presbyterian College. Four are school teachers: one from Richland School District One, two from Lexington School District One, and one from Florence School District Four. The remaining seven are a representative from the HEM, two representatives from Pearson Early Learning, and four representatives from the SDE. Advisory committee meetings were held November 2001, January 2002, March 2002, and April 2002.

November 2001 Meeting

As stated earlier, the November meeting was the SCRA advisory committee's initial one. Committee members discussed the following topics:

- current student home reports,
- consistency of teacher ratings,
- removing the "progress" rating column from the checklist and the student home report,
- changing the term *proficient* so as not to lead to a comparison to the PACT,
- statewide reporting of SCRA results,
- using electronic checklists, and
- staff development for summer 2002 to train lead teachers for each school.

Three subcommittees were formed to address these three key issues for 2002–03: guidelines revisions, report revisions, and general procedures.

January 2002 Meeting

Each of the subcommittees had at least one facilitator with expertise in that particular area. Two representatives from Pearson Early Learning served in this role for the subcommittee on guidelines revisions. A representative from the HEM assisted with the report revision subcommittee, and a representative from the SDE worked with the general procedures subcommittee.

At the meeting, a representative from SERVE presented a general overview of the procedures and findings of the SERVE review of the SCRA. Draft copies of the teacher survey were provided to committee members. Other topics discussed were definition of ratings, possible approaches to the proposed lead teacher training for summer 2002, and the use of an SCRA alternate assessment.

March 2002 Meeting

The committee reviewed the new guidelines draft from Pearson Early Learning. There was discussion on the inclusion of the state curriculum standards in the guidelines. The committee also discussed the number of student ratings with the SCRA per year and ways to communicate to teachers the changes for 2002–03.

April 2002 Meeting

The advisory committee reviewed the draft copies of the score reports, checklists, and guidelines. The score reports and the checklists contained changes in the wording of the ratings and the rating periods, and the “progress” rating column had been removed. Two versions of the student report (one for teachers and one for parents) had also been developed. Additionally, functional components had been changed to make them consistent with the state curriculum standards in English language arts.

Committee members agreed to have four ratings per year, with only two of them required: “Performance to date (Winter)” and “Year-End Performance.” Pearson Early Learning representatives presented notes on a two-day summer workshop agenda for teachers and administrators. Teachers and administrators were scheduled to attend the first day; the second day was designed for teachers only.

1.3 REVISIONS FOR 2002–03

As is stated above, the advisory committee meetings during 2001–02 resulted in revisions to the SCRA. At its initial meeting in November, the committee discussed redesigning the student home report, removing the “progress” part of the checklist, changing the timing of the data collection and reports, and writing new guidelines documents with more specific ties to the state standards.

Developmental Guidelines

The new guidelines reflect the language of the fourth version of the WSS, the 2000 South Carolina mathematics curriculum standards, and the 2002 South Carolina English language arts curriculum standards. The guidelines are issued as one document that covers both kindergarten and grade one. For each indicator, the left-hand page contains the guidelines for kindergarten, and the right-hand page contains the guidelines for grade one. On each page of the guidelines, the SCRA indicator, the related state standards for English language arts and mathematics, the rationale, and examples are given. The guidelines also include an introduction, an explanation of the ratings, and acknowledgements.

Checklists

Changes in the checklists involved the subject domains titles, the functional components, the performance indicators, and the assessment periods.

There were two changes in the titles of the subject domains: the “Language and Literacy” domain was renamed “English Language Arts,” and the “Mathematical Thinking” domain was renamed “Mathematics.” The title “Personal and Social Development” remained the same.

Table 1.1 reflects the changes in the functional components. The number of functional components was decreased from fifteen in 2001–02 to fourteen for 2002–03. The two functional components “Listening” and “Speaking” were combined under the new functional component labeled “Communication.”

Under “Personal and Social Development” domain, the number of indicators for kindergarten and grade one in 2002–03 were decreased to thirteen, compared to sixteen in 2001–02. In the kindergarten “English Language Arts” domain, the number of indicators for 2002–03 was increased to twelve, from eleven in 2001–02. The number of first-grade “English Language Arts” indicators remained at twelve. In the kindergarten “Mathematics” domain, the number of indicators for 2002–03 was increased to fourteen, from thirteen in 2001–02. For the first grade, the number of indicators for “Mathematics” remained at fourteen. After these changes were made, the SCRA development checklists for 2002–03 were comprised of thirty-seven performance indicators for kindergarten and thirty-nine performance indicators for the first grade; and both grade levels had the same functional components for each of the three domains as shown below in table 1.1.

TABLE 1.1

2002–03 SCRA Functional Components for Kindergarten and Grade One

Domain	Functional Component
Personal and Social Development	A. Self-Concept
	B. Self-Control
	C. Approaches to Learning
	D. Interaction with Others
	E. Social Problem-Solving
English Language Arts	A. Communication
	B. Reading
	C. Writing
Mathematics	A. Mathematical Processes
	B. Number and Operations
	C. Patterns, Relationships, and Functions
	D. Geometry and Spatial Relations
	E. Measurement
	F. Data Collection and Probability

The SCRA mastery levels *not yet*, *in process*, and *proficient* were changed to the performance indicators *rarely or never demonstrates*, *sometimes demonstrates*, and *consistently demonstrates*. The descriptions of the assessment periods were changed from “Fall,” “Winter,” and “Spring,” with the overall rating in “Spring,” to the following: “Performance to date (Fall),” “Performance to date (Winter),” “Performance to date (Spring),” and “Year-End Performance.”

Reports

Changes were also made in the student home report. On the new student home report, all performance indicators will appear. However, there will be two versions of the report, one for parents and one for teachers. The parent version will show the indicators but will have ratings only for the functional components, whereas the teacher version will show the ratings for both the indicators and the functional components. The progress indicators will no longer be on the report.

Chapter 2

TEST ADMINISTRATION AND SCORING

2.1 OVERVIEW

Using machine-scannable paper copies of the checklist, in 2002–03 a teacher made four assessments of his or her students during the school year (fall, winter, spring, and end of year) to determine their current level of achievement—*consistently demonstrates*, *sometimes demonstrates*, or *rarely or never demonstrates*—with regard to each of the indicators. The machine-scannable checklists were sent to the HEM San Antonio scoring center for scoring at the end of the school year.

All students enrolled in kindergarten and grade one for at least one month were assessed. Provisions were made for those students who transferred either into the school or out to another school. For each kindergarten and first-grade student new to the school, a scannable checklist was to be filled out by the receiving teacher if the student arrived within four weeks before the end of the assessment period. The transferring school was required to forward the scannable checklist to the new school as part of the student’s permanent file. However, the transferring school was not required to send the teacher copy of the checklist or the teacher file of observations and student work.

In lieu of completing paper checklists, school districts that agreed to participate in the online version of the SCRA (see section 2.3) could submit their data electronically. Approximately 25 percent of kindergarten and first-grade teachers used the South Carolina Readiness Assessment Profile Interactive (SCRAPI) online system in the 2002–03 pilot. The online system uploads student ratings to a state database so that results are always current and no physical reports need to be submitted.

2.2 ADMINISTRATOR TRAINING

After the 2000 and 2001 training workshops, the trained district personnel became responsible for training new teachers. In addition, approximately ten thousand copies of the *South Carolina Readiness Assessment Kindergarten and First Grade Developmental Guidelines* (Rebus and SDE 2002) were distributed to teachers statewide in July and August 2002.

The guidelines prior to 2002–03 were rewritten to reflect changes in the WSS and the revised South Carolina English language arts curriculum standards. The new 2002–03 guidelines also combined kindergarten and the first grade in a single document for easy reference. Additionally, the standards associated with each indicator in mathematics and English language arts were listed beneath the indicator along with the rationale and examples.

In September 2002, district personnel who participated in the pilot of the SCRAPI online system were trained in the use of the system in a workshop conducted by the SDE. In addition, the

training was made available during a live broadcast over the Instructional Television network on October 6, 2002.

In training workshops conducted by the SDE in March 2003, district test coordinators received copies of the *South Carolina Readiness Assessment Developmental Checklists District and School Test Administrator Manual 2002–2003* (HEM 2002). In April 2003, schools received their copies of the manual. Although the manual contains procedures for returning paper checklists, it also gives explanations for the student demographic coding that must be completed for both the paper and the online systems.

2.3 SOUTH CAROLINA READINESS ASSESSMENT PROFILE INTERACTIVE

The South Carolina Readiness Assessment Profile Interactive (SCRAPI) online system allows teachers to submit their completed checklists electronically. This section summarizes the findings of Porchea and Casteel (2003) regarding the system.

Initial implementation of the SCRAPI system during the 2002–03 school year was invaluable in generating feedback about the system. The feedback was used to enhance the user-friendliness and utility of the system for subsequent SCRA administration years.

History

Since the inception of the SCRA, teachers had been concerned about the amount of paperwork involved, and efforts were made to ease their concerns. The SCRA checklists are now shorter than the WSS checklists, from which they were derived. With advisory committee support, the SDE eased the burden in 2001–02 and 2002–03 by reducing the mandatory rating periods from three (fall, winter, and end of year) to two (winter and end of year) and providing preprinted labels to override the demographic page, which would have to be completed manually if they did not have the preprinted labels.

In 2001, the SDE contracted with the state's Horry County School District, which had already begun creating an online reporting system for its own use. The system was first tested by all 2001–02 kindergarten and first-grade teachers in the Horry County School District and in Dillon School District Two. Approximately 250 teachers completed the ratings on over 5,100 students.

Once the SCRAPI software was installed in the state computer system, student demographic data were loaded into the SCRAPI system. Districts submitted student, teacher, and parent information to the state at the beginning of the 2002–03 school year. The data were then loaded into the SCRAPI system. Hence, teachers logging on to the system for the first time found their classroom information already on the system.

District testing coordinators and technology coordinators were invited to volunteer their districts, as a whole or in part, to pilot SCRAPI in the 2002–03 school year. Twenty-two districts successfully participated in the pilot. Some of these districts used SCRAPI districtwide; others used SCRAPI in only some of their schools.

Online System Features

Since the teacher and student information was preloaded into SCRAPI, teachers logging on to the system for the first time enter minimal identifying information and an e-mail address. The system assigns a random password and sends it to the teacher via e-mail.

Once logged on to the system, the teachers have access to several menu options. Teachers complete online student checklists, which are designed to resemble paper checklists.

Teachers can also view a class summary showing the proportion of each student's checklist that has been completed as well as the proportion of the total class with completed checklists. Since the summary pages in SCRAPI access the state database, the summaries are always current. A summary of the total ratings is available (*rarely or never demonstrates, sometimes demonstrates, or consistently demonstrates*) for teachers to be able to assess their overall class levels at any given point in the year.

School administrators can view a teacher-ratings summary that shows where teachers are in the rating process (i.e., the percentage of checklists completed). School administrators can view individual student ratings but do not have privileges to add or modify student ratings.

District administrators can view a summary that shows the number of ratings completed at each of their schools. (A district administrator can also function as a school administrator in any of the schools in his or her district.) State administrators can view a summary of all the districts that shows the number of ratings completed in each district as well as the total number of students in the system. (A state administrator can also function as a district/school administrator for any district/school.)

Transferring Students

A common complaint made by school personnel had been that students transfer into their schools during the school year without all of the necessary documents from the previous school's files. Even when the sending school included all documents from the student's permanent records, often the SCRA checklist was not included because it may have been stored in the classroom rather than in the school office. Now, however, if a student is in the SCRAPI system, his or her ratings are already electronically stored and it is simply a matter of reassigning the student to the receiving school.

School administrators handle student transfers. SCRAPI has a feature that allows school administrators to search for students with a similar name or other identifying information. Once the student record is found and selected, if the record is marked *inactive* (meaning the student is not currently assigned to a school), the school administrator selects the student record and designates it to the new teacher. If the student record is marked *active*, the SCRAPI system automatically sends an e-mail to the school administrator at the sending school requesting that the student record be released. Once the sending school releases it, another e-mail is generated to notify the receiving school that the student record can now be transferred.

2.4 SCORING

At the end of the 2002–03 school year, the data collected through SCRAPI were sent to the HEM for processing. HEM consolidated the SCRAPI data with the SCRA data collected from paper checklists. Scoring of the 2002–03 checklists for kindergarten and grade one proceeded in the same manner as in previous years. Under contract to the SDE, the HEM was responsible for receiving, scanning, scoring, and reporting SCRA results. The HEM and the SDE determined the scoring algorithms and report formats for the readiness assessment. Consultants from Rebus, the HEM, and the SDE determined the format and content of score reports for students, districts, and the state.

Using the checklists, teachers made four assessments of their students' performance with regard to the curriculum standards: performance to date (fall), performance to date (winter), performance to date (spring), and year-end performance. Only year-end performance ratings were used in the reporting process. For paper checklists, number-two lead pencils were to be used to mark all ratings on the official copy. Schools then sent the completed scannable checklists to the district testing coordinator, who in turn sent them to the HEM San Antonio scoring center for scoring at the end of the school year.

The HEM's technical proposal to the SDE on the readiness assessment for first and second graders explains the quality control and scoring systems:

Harcourt's WorkFlow System is designed to provide several checks/edits of scannable documents for accuracy. The WorkFlow System allows for a second edit to check for: 1) the completeness of information provided on scannable materials; 2) scannable documents returned in poor condition; and 3) missing students', schools' or districts' documents. Several workstations are involved in this process.

The workstation executes the resolve/edit program and produces the edit list. The edit list is a sequential listing of error suspect records, including SSIDs. The suspect field(s) within a record are printed with a flag. This flag means that the data field did not pass our edit requirements. . . . In each case, an error suspect record must be reviewed by an editor who references the source answer document to ensure that the integrity of the data provided is maintained. Sequence numbers of the answer documents aid the editor in this review. . . .

Score reports for the South Carolina Readiness Assessment will be printed only after an internal quality assurance and an external Department quality assurance check have verified accuracy. To further ensure accurate reporting for all districts, Harcourt will conduct a pre-print and post-print quality assurance check of every school, district, and building report. (HEM 1999, 45–46)

2.5 SCORE REPORTING

Score reporting for the 2002–03 kindergarten and grade-one data was the same as that for previous years. After the data were collected and other data analyses had been conducted, the individual student home reports; the class rosters; and the school, district, and state summaries were produced and shipped to the school districts. The district test coordinator was responsible

for ensuring that schools received the student home reports and other reports at the beginning of the 2003–04 school year.

Student Home Report

The 2002–03 student home report was a single 8½ x 11 inch page that used the performance levels *rarely or never demonstrates*, *sometimes demonstrates*, and *consistently demonstrates* to designate the student’s strengths and needs in each of the functional components. Explanations of the performance levels and performance indicators were also provided in the student home report. The component rating was reported using an algorithm based on the scoring rules described in table 2.1 on the next page.

Class Roster

The class roster was an 8½ x 11 inch multipage document provided for every teacher by grade level. Students’ names were listed in alphabetical order, and each student’s performance was reported for each functional component.

School, District, and State Summaries

The school report was titled “Class List with School Summary”; the district report, “School List with District Summary,” and the state report, “District List with State Summary.” These 8½ x 11 inch multipage reports provided summary information for three subject domains called “Personal and Social Development,” “English Language Arts,” and “Mathematics.” Performance-level information was reported by the number and the percentage of students at the *rarely or never demonstrates*, *sometimes demonstrates*, and *consistently demonstrates* levels for each of the functional components within the three domains. The last page of the report provided a summary of demographic data used on the checklists.

TABLE 2.1

Frequency of Performance Ratings Needed to Achieve Various Component Ratings

Number of Component Ratings	Frequency of Performance Ratings			Component Rating	
	Consistently Demonstrates	Sometimes Demonstrates	Rarely or Never Demonstrates	Title	Score
1	1	0	0	consistently demonstrates	3
1	0	1	0	sometimes demonstrates	2
1	0	0	1	rarely or never demonstrates	1
2	2	0	0	consistently demonstrates	3
2	1	1	0	sometimes demonstrates	2
2	1	0	1	sometimes demonstrates	2
2	0	2	0	sometimes demonstrates	2
2	0	1	1	sometimes demonstrates	2
2	0	0	2	rarely or never demonstrates	1
3	3	0	0	consistently demonstrates	3
3	2	1	0	consistently demonstrates	3
3	2	0	1	sometimes demonstrates	2
3	1	2	0	sometimes demonstrates	2
3	1	1	1	sometimes demonstrates	2
3	1	0	2	sometimes demonstrates	2
3	0	3	0	sometimes demonstrates	2
3	0	2	1	sometimes demonstrates	2
3	0	1	2	rarely or never demonstrates	1
3	0	0	3	rarely or never demonstrates	1
4	4	0	0	consistently demonstrates	3
4	3	1	0	consistently demonstrates	3
4	3	0	1	sometimes demonstrates	2
4	2	2	0	sometimes demonstrates	2
4	2	1	1	sometimes demonstrates	2
4	2	0	2	sometimes demonstrates	2
4	1	3	0	sometimes demonstrates	2
4	1	2	1	sometimes demonstrates	2
4	1	1	2	sometimes demonstrates	2
4	1	0	3	sometimes demonstrates	2
4	0	4	0	sometimes demonstrates	2
4	0	3	1	sometimes demonstrates	2
4	0	2	2	sometimes demonstrates	2
4	0	1	3	rarely or never demonstrates	1
4	0	0	4	rarely or never demonstrates	1
5	5	0	0	consistently demonstrates	3
5	4	1	0	consistently demonstrates	3
5	4	0	1	sometimes demonstrates	2
5	3	2	0	sometimes demonstrates	2
5	3	1	1	sometimes demonstrates	2
5	3	0	2	sometimes demonstrates	2
5	2	3	0	sometimes demonstrates	2
5	2	2	1	sometimes demonstrates	2
5	2	1	2	sometimes demonstrates	2
5	2	0	3	sometimes demonstrates	2
5	1	4	0	sometimes demonstrates	2
5	1	3	1	sometimes demonstrates	2
5	1	2	2	sometimes demonstrates	2
5	1	1	3	rarely or never demonstrates	1
5	1	0	4	rarely or never demonstrates	1
5	0	5	0	sometimes demonstrates	2
5	0	4	1	sometimes demonstrates	2
5	0	3	2	sometimes demonstrates	2
5	0	2	3	rarely or never demonstrates	1
5	0	1	4	rarely or never demonstrates	1
5	0	0	5	rarely or never demonstrates	1

Chapter 3

TECHNICAL CHARACTERISTICS OF THE INDICATORS AND RATINGS

As noted previously, the SCRA development checklists for 2002–03 were comprised of thirty-seven performance indicators for kindergarten and thirty-nine performance indicators for the first grade. Values for each indicator were assigned four times during the year: performance to date (fall), performance to date (winter), performance to date (spring), and year-end performance. Only the winter and year-end ratings were required. Values for the fourteen functional component ratings were derived from the year-end performance ratings. This section documents the technical characteristics of the indicators and ratings. For ease of terminology, these quantities are referred to globally as “items” in the remaining part of this chapter.

For each item, the SDE computed statistics such as the number of students, item mean, and item/total score correlation. Correlations were computed between the item and the total score (which includes that item). Appendix A provides these item statistics in detail for the year-end performance indicators and the year-end performance ratings in kindergarten and the first grade. The 2002–03 statistics in this chapter include data from 212 students (of whom 106 were kindergarteners and 106 were first graders) who received alternate scoring as reported by the schools. Students who receive alternate scoring on the SCRA are rated only on the indicators that align with their individualized education plan (IEP) goals and objectives.

3.1 YEAR-END PERFORMANCE INDICATORS

Table 3.1 gives a summary of the year-end performance indicators in statewide kindergarten and the first grade as well as for each gender (male and female) and ethnicity (African American and white) group. The effective sample size for each statistic varied since not all students had complete data.

TABLE 3.1
Summary of Major Statistics for Year-End Performance Indicators

	Student Group				
	All Students	Males	Females	African Americans	Whites
Kindergarten					
Number of students	46,704	24,312	22,175	17,201	24,503
Number of items	37	37	37	37	37
Average of item means	2.698	2.640	2.762	2.612	2.768
Average of item standard deviations	0.523	0.559	0.470	0.568	0.469
Average of correlations between items and total score	0.743	0.746	0.730	0.744	0.731
Alpha coefficient	0.979	0.979	0.978	0.979	0.977
Grade 1					
Number of students	48,652	25,074	23,476	18,111	26,287
Number of items	39	39	39	39	39
Average of item means	2.669	2.614	2.729	2.561	2.750
Average of item standard deviations	0.543	0.574	0.497	0.595	0.483
Average of correlations between items and total score	0.750	0.752	0.742	0.750	0.733
Alpha coefficient	0.980	0.980	0.980	0.980	0.978

Source: State Department of Education

3.2 YEAR-END PERFORMANCE RATINGS

Table 3.2 provides a summary of the year-end performance ratings in statewide kindergarten and the first grade as well as for each gender (male and female) and ethnicity (African American and white) group. The effective sample size for each statistic varied since not all students had complete data. Tables 3.3 and 3.4 provide the state summary since 2000–01 and illustrate the changes in performance ratings terminology since 2000–01.

TABLE 3.2

Summary of Major Statistics for Year-End Performance Ratings

	Student Group				
	All Students	Males	Females	African Americans	Whites
Kindergarten					
Number of students	46,674	24,298	22,160	17,191	24,488
Number of items	14	14	14	14	14
Average of item means	2.675	2.618	2.738	2.583	2.751
Average of item standard deviations	0.519	0.551	0.473	0.561	0.469
Average of correlations between items and total score	0.763	0.767	0.750	0.761	0.752
Alpha coefficient	0.950	0.951	0.947	0.949	0.947
Grade 1					
Number of students	48,619	25,056	23,461	18,091	26,274
Number of items	14	14	14	14	14
Average of item means	2.649	2.595	2.708	2.534	2.735
Average of item standard deviations	0.535	0.562	0.497	0.582	0.480
Average of correlations between items and total score	0.769	0.771	0.761	0.767	0.753
Alpha coefficient	0.952	0.952	0.950	0.951	0.947

Source: State Department of Education

TABLE 3.3

State Summary of Year-End Performance Ratings of the Kindergarten Assessment for First-Grade Readiness

2000–01 and 2001–02 Terminology	2000–01			2001–02			2002–03 Terminology	2002–03		
	Not Yet	In Process	Proficient	Not Yet	In Process	Proficient		Rarely or Never Demonstrates	Sometimes Demonstrates	Consistently Demonstrates
Language and Literacy							English Language Arts			
Listening	2%	39%	59%	2%	37%	60%	Communication	2%	32%	66%
Speaking	2%	34%	64%	2%	32%	65%	Reading	3%	27%	70%
Literature and Reading	3%	33%	64%	3%	33%	64%	Writing	3%	21%	76%
Writing	2%	38%	60%	2%	37%	61%				
Mathematical Thinking							Mathematics			
Approach to Mathematical Thinking	3%	41%	55%	3%	39%	58%	Mathematical Processes	6%	35%	59%
Number Concept and Operations	2%	32%	66%	2%	30%	68%	Number and Operations	2%	24%	74%
Patterns and Relationships	2%	24%	74%	2%	23%	75%	Patterns, Relationships, and Functions	1%	21%	78%
Geometry and Spatial Relationships	1%	32%	67%	1%	30%	69%	Geometry and Spatial Relations	1%	24%	75%
Measurement	4%	36%	60%	4%	34%	63%	Measurement	4%	29%	67%
Probability and Statistics	6%	42%	52%	5%	38%	57%	Data Collection and Probability	5%	31%	64%
Personal and Social Development							Personal and Social Development			
Self-Concept	2%	38%	60%	2%	37%	62%	Self-Concept	2%	36%	62%
Self-Control	2%	26%	72%	2%	24%	74%	Self-Control	2%	20%	78%
Approach to Learning	2%	36%	62%	2%	34%	64%	Approaches to Learning	5%	29%	67%
Interactions with Others	2%	33%	65%	2%	32%	67%	Interaction with Others	1%	21%	79%
Conflict Resolution	2%	40%	58%	2%	37%	61%	Social Problem-Solving	3%	24%	73%

Note: For each school year, the sum of the percentages may equal 99 percent or 101 percent due to rounding to the nearest whole percentage.

TABLE 3.4

**State Summary of Year-End Performance Ratings of the
First-Grade Assessment for Second-Grade Readiness**

2000–01 and 2001–02 Terminology	2000–01			2001–02			2002–03 Terminology	2002–03		
	Not Yet	In Process	Pro- ficient	Not Yet	In Process	Pro- ficient		Rarely or Never Demonstrates	Sometimes Demonstrates	Consistently Demonstrates
Language and Literacy							English Language Arts			
Listening	3%	43%	54%	3%	41%	56%	Communication	3%	35%	62%
Speaking	3%	39%	58%	3%	36%	61%				
Literature and Reading	4%	37%	59%	4%	35%	61%	Reading	4%	29%	67%
Writing	4%	43%	53%	4%	41%	55%	Writing	6%	29%	65%
Mathematical Thinking							Mathematics			
Approach to Mathematical Thinking	3%	44%	53%	3%	41%	56%	Mathematical Processes	3%	33%	64%
Number Concept and Operations	2%	39%	58%	2%	37%	60%	Number and Operations	3%	21%	75%
Patterns and Relationships	1%	30%	69%	1%	27%	72%	Patterns, Relationships, and Functions	1%	20%	78%
Geometry and Spatial Relationships	2%	42%	56%	2%	39%	60%	Geometry and Spatial Relations	2%	31%	67%
Measurement	2%	44%	54%	2%	42%	56%	Measurement	3%	27%	70%
Probability and Statistics	3%	38%	59%	3%	36%	62%	Data Collection and Probability	4%	34%	63%
Personal and Social Development							Personal and Social Development			
Self-Concept	3%	41%	57%	2%	39%	58%	Self-Concept	3%	38%	59%
Self-Control	3%	28%	70%	3%	26%	71%	Self-Control	2%	21%	77%
Approach to Learning	3%	37%	61%	2%	35%	63%	Approaches to Learning	5%	27%	68%
Interactions with Others	2%	34%	64%	2%	32%	66%	Interaction with Others	1%	23%	76%
Conflict Resolution	3%	47%	50%	3%	43%	54%	Social Problem-Solving	5%	30%	64%

Note: For each school year, the sum of the percentages may equal 99 percent or 101 percent due to rounding to the nearest whole percentage.

Chapter 4

VALIDITY

This chapter reports on measuring fairness using differential item functioning (DIF) for categories of gender and ethnicity.

4.1 OVERALL VIEW OF DIF ANALYSIS

General Description

One threat to the validity of an assessment is test bias, or the unfair advantage of one group over another on the test. DIF occurs, in the words of Clauser and Mazor, when “examinees from different groups have differing probabilities or likelihoods of success on an item, after they have been matched” on a characteristic of interest such as ability or achievement on the test as a whole. “DIF is a necessary but not sufficient condition for item bias,” they explain (Clauser and Mazor 1998, 31). DIF focuses on *item* validity as opposed to *test* validity because the procedures assume that the test as a whole is a “good” measure.

For documentation purposes, SCRA items (performance indicators and performance ratings) were subjected to a formal DIF analysis based on the Mantel-Haenszel (MH) procedure, which has a long tradition in DIF analysis and is considered effective and efficient (see Clauser and Mazor 1998, Hills 1989). The MH procedure uses both a statistical significance test and an analysis of the effect size.

In the MH procedure, total scores (on the entire scale) are typically used to group students into strata. Students in each stratum were considered as equivalent in terms of the construct under assessment. Then, for each item, the students in the “focal” and “reference” groups were compared on the basis of their mean score on the item. The term *focal* refers to the group of interest for DIF—in this case, female or African American. The comparison or reference group was male or white, depending upon whether the DIF analysis was for gender or ethnicity.

In traditional National Assessment of Educational Progress (NAEP) DIF analysis for multiple-choice items, items are classified as “A,” “B,” or “C” in terms of DIF. “A” items are considered to be free of DIF. “B” items may be used, but if there is a choice among otherwise equivalent items, it is considered desirable to select for inclusion in a test those items with the smallest absolute value of MH DIF. “C” items are to be selected only if they are essential to meet test specifications (Zwick and Ercikan 1989, 58–59). For other items with more than two ordered response categories (such as SCRA items), the corresponding NAEP DIF classifications are “AA,” “BB,” and “CC.”

Standardized Mean Difference Procedure

For the 2002–03 year-end performance SCRA data, the MH procedure for polytomous items was used. This procedure is based on the Cochran-Mantel-Haenszel (CMH) procedure for ordered categorical variables and an analysis of the standardized mean difference (SMD). (See Dorans, Schmitt, and Bleistein 1992; Zwick and Thayer 1996; Zwick, Thayer, and Mazzeo 1997.) Values of the SMD serve as measures of effect size for constructed-response items.

The MH procedure was implemented using SAS (Stokes, Davis, and Koch 2001). Students were grouped into ten strata on the basis of their complete total raw scores. The FREQ procedure provided the MH statistic with a p -value. The standard deviation (SD) of the combined group, the unweighted mean (MF) of the focal group, and the weighted mean (MR) of the reference group were computed using SAS. The MR was obtained by taking the mean of the reference group in each stratum and weighting it according to the number of focal students in the same stratum. The MR subtracted from the MF is the SMD. An estimate of the effect size (ES) of the mean was then obtained by dividing the SMD by the SD. Following the recommendation of the NAEP (see Allen, Carlson, and Zalanak 1999), the combined group SD was used as a scale measure for the item. With several DIF analyses to be carried out for the same set of data, it was necessary to keep the scale constant so that all ESs *for the same item* would be expressed on a common scale and therefore compatible with each other across different DIF analyses.

NAEP DIF Classification

The following ETS/NAEP rules were used to classify items as “AA,” “BB,” and “CC.”

Rule 1: If the p -value > 0.05 , then classify the item as AA.

Otherwise, use the following rules based on the absolute value of the ES:

Rule 2: If $\text{abs}(\text{ES})$ is less than or equal to 0.17, then classify the item as AA.

Rule 3: If $\text{abs}(\text{ES})$ is greater than 0.17 and less than or equal to 0.25, then classify the item as BB.

Rule 4: If $\text{abs}(\text{ES})$ is greater than 0.25, then classify the item as CC.

4.2 RESULTS OF DIF ANALYSES

Table 4.1 provides a summary of DIF classifications for the year-end performance indicators for both kindergarten and grade-one students. Similar data for the year-end performance ratings are listed in table 4.2. For the year-end performance indicators, the total 76 indicators were classified as “AA” for 74 cases and “BB” for 2 cases in gender DIF. All 76 cases were classified as “AA” in ethnicity DIF. For the year-end performance ratings, the total 28 ratings were classified as “AA” for 27 cases and “BB” for 1 case in gender DIF. All 28 ratings were classified as “AA” in ethnicity DIF. The data clearly indicate that the checklists are reasonably free of gender and ethnicity DIF. Appendix B provides complete DIF data for year-end performance indicators and year-end performance ratings. The DIF analyses also include data from students who received alternate scoring as reported by the schools.

TABLE 4.1

Summary of DIF Classification for Year-End Performance Indicators

Reference Group	Focal Group	Total Number of Items	NAEP DIF Classification		
			AA	BB	CC
Kindergarten					
Males	Females	37	36	1	---
Whites	African Americans	37	37	---	---
First Grade					
Males	Females	39	38	1	---
Whites	African Americans	39	39	---	---
Both Grades					
Males	Females	76	74	2	---
Whites	African Americans	76	76	---	---

TABLE 4.2

Summary of DIF Classification for Year-End Performance Ratings

Reference Group	Focal Group	Total Number of Items	NAEP DIF Classification		
			AA	BB	CC
Kindergarten					
Males	Females	14	14	---	---
Whites	African Americans	14	14	---	---
First Grade					
Males	Females	14	13	1	---
Whites	African Americans	14	14	---	---
Both Grades					
Males	Females	28	27	1	---
Whites	African Americans	28	28	---	---

Appendix A

ITEM STATISTICS FOR YEAR-END PERFORMANCE INDICATORS AND YEAR-END PERFORMANCE RATINGS

Provided here are the following tables with detailed DIF analysis results for gender and ethnicity in kindergarten and the first grade:

Table A1: Item Statistics: Year-End Performance Indicators for Kindergarten and First Grade

Table A2: Item Statistics: Year-End Performance Ratings for Kindergarten and First Grade

The following terms are used in these two tables:

Indicator	=	item sequence as it appears in the checklist
Functional Component	=	functional component name
Item Mean	=	arithmetic mean of the item (Responses were assigned the following weights: 1 for <i>rarely or never demonstrates</i> , 2 for <i>sometimes demonstrates</i> , and 3 for <i>consistently demonstrates</i> . The mean includes all students with a rating on the item regardless of whether they have complete ratings on all other items.)
Item-Scale Correlation	=	correlation between the item and the total score on the scale (Only records with responses to all the checklist items were included in the correlation. The term <i>scale</i> denotes the total score of all items that form the checklist.)
N	=	number of students with responses to the item

TABLE A.1**Item Statistics: Year-End Performance Indicators for Kindergarten and First Grade**

Kindergarten				First Grade			
Indicator	Item Mean	N	Item-Scale Correlation	Indicator	Item Mean	N	Item-Scale Correlation
1	2.670	46,498	0.695	1	2.639	48,464	0.725
2	2.621	46,485	0.761	2	2.581	48,431	0.781
3	2.677	46,513	0.554	3	2.659	48,479	0.573
4	2.806	46,398	0.604	4	2.787	48,383	0.628
5	2.751	46,505	0.641	5	2.728	48,462	0.669
6	2.725	46,483	0.755	6	2.698	48,469	0.764
7	2.572	46,446	0.765	7	2.588	48,429	0.758
8	2.573	46,483	0.794	8	2.560	48,457	0.811
9	2.788	46,515	0.589	9	2.748	48,451	0.604
10	2.837	46,461	0.596	10	2.779	48,414	0.616
11	2.768	46,427	0.690	11	2.750	48,387	0.704
12	2.759	46,441	0.604	12	2.744	48,404	0.606
13	2.704	46,477	0.668	13	2.591	48,431	0.733
14	2.682	46,326	0.778	14	2.657	48,328	0.777
15	2.637	46,243	0.785	15	2.586	48,236	0.805
16	2.677	46,208	0.749	16	2.646	48,173	0.760
17	2.638	46,236	0.785	17	2.592	48,208	0.794
18	2.780	46,206	0.768	18	2.773	48,239	0.725
19	2.758	46,188	0.804	19	2.783	48,197	0.769
20	2.673	46,209	0.790	20	2.664	48,220	0.773
21	2.660	46,244	0.778	21	2.597	48,188	0.785
22	2.705	46,218	0.805	22	2.647	48,239	0.799
23	2.730	46,175	0.801	23	2.568	48,204	0.819
24	2.713	46,158	0.785	24	2.595	48,191	0.780
25	2.702	46,269	0.820	25	2.562	48,310	0.809
26	2.550	46,472	0.811	26	2.694	48,458	0.773
27	2.566	46,417	0.817	27	2.613	48,378	0.803
28	2.756	46,410	0.780	28	2.769	48,388	0.763
29	2.721	46,410	0.803	29	2.721	48,328	0.790
30	2.789	46,390	0.769	30	2.542	48,412	0.790
31	2.802	46,431	0.760	31	2.791	48,337	0.765
32	2.790	46,352	0.764	32	2.803	48,401	0.751
33	2.759	46,443	0.772	33	2.774	48,337	0.768
34	2.663	46,419	0.778	34	2.635	48,397	0.783
35	2.567	46,403	0.765	35	2.631	48,407	0.772
36	2.664	46,387	0.767	36	2.665	48,376	0.771
37	2.580	46,450	0.754	37	2.684	48,455	0.760
				38	2.663	48,395	0.798
				39	2.593	48,415	0.792

TABLE A.2**Item Statistics: Year-End Performance Ratings for Kindergarten and First Grade**

Functional Component	Kindergarten			First Grade		
	Item Mean	N	Item-Scale Correlation	Item Mean	N	Item-Scale Correlation
Self-Concept	2.599	46,375	0.717	2.566	48,336	0.740
Self-Control	2.764	46,201	0.627	2.748	48,197	0.655
Approaches to Learning	2.620	46,218	0.808	2.626	48,230	0.821
Interaction with Others	2.778	46,065	0.674	2.745	48,049	0.678
Social Problem-Solving	2.704	46,477	0.668	2.591	48,431	0.733
Communication	2.634	45,614	0.813	2.594	47,631	0.812
Reading	2.667	45,278	0.831	2.631	47,388	0.811
Writing	2.726	45,655	0.833	2.583	47,817	0.828
Mathematical Processes	2.534	46,340	0.802	2.613	48,317	0.779
Number and Operations	2.713	46,254	0.792	2.720	48,127	0.806
Patterns, Relationships, and Functions	2.765	46,258	0.780	2.771	48,228	0.769
Geometry and Spatial Relations	2.733	46,221	0.780	2.649	48,223	0.760
Measurement	2.636	46,151	0.802	2.665	48,236	0.790
Data Collection and Probability	2.580	46,450	0.754	2.589	48,338	0.780

Appendix B

DIFFERENTIAL ITEM FUNCTIONING FOR YEAR-END PERFORMANCE INDICATORS AND YEAR-END PERFORMANCE RATINGS

Provided here are the following tables with detailed DIF analysis results for gender and ethnicity in kindergarten and the first grade:

Table B.1	DIF Data: Year-End Performance Indicators for Kindergarten and Gender
Table B.2	DIF Data: Year-End Performance Indicators for Kindergarten and Ethnicity
Table B.3	DIF Data: Year-End Performance Indicators for First Grade and Gender
Table B.4	DIF Data: Year-End Performance Indicators for First Grade and Ethnicity
Table B.5	DIF Data: Year-End Performance Ratings for Kindergarten and Gender
Table B.6	DIF Data: Year-End Performance Ratings for Kindergarten and Ethnicity
Table B.7	DIF Data: Year-End Performance Ratings for First Grade and Gender
Table B.8	DIF Data: Year-End Performance Ratings for First Grade and Ethnicity

The following terms are used in the tables:

Indicator	= item sequence as it appears in the checklist
Functional Component	= functional component name
<i>p</i> -value	= MH statistic's probability for statistical significance (With a large number of students in the DIF groups, the <i>p</i> -value can be very small and may be indicated by 0.000, which means the probability is less than 0.0005.)
Focal Mean	= unweighted mean of the focal group
Reference Mean	= weighted mean of the reference group where weights are the proportion of focal students in each stratum over all the students in the focal group
SD	= standard deviation of the item/rating over all focal and reference students
Effect Size	= the SMD (the focal mean minus the reference mean) divided by the standard deviation
DIF	= classification based on NAEP rules for differential item functioning [plus sign (+) = in favor of the focal group; minus sign (-) = against the focal group]

TABLE B.1**DIF Data: Year-End Performance Indicators for Kindergarten and Gender**

Indicator	<i>P</i>-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
1	0.254	2.732	2.738	0.522	-0.011	AA
2	0.000	2.713	2.686	0.571	0.049	AA
3	0.000	2.789	2.692	0.526	0.185	BB+
4	0.000	2.890	2.823	0.429	0.157	AA
5	0.000	2.836	2.781	0.483	0.114	AA
6	0.019	2.792	2.787	0.501	0.008	AA
7	0.000	2.686	2.638	0.605	0.080	AA
8	0.002	2.666	2.656	0.602	0.016	AA
9	0.000	2.842	2.829	0.444	0.029	AA
10	0.008	2.881	2.872	0.394	0.021	AA
11	0.000	2.832	2.815	0.466	0.038	AA
12	0.000	2.832	2.793	0.473	0.082	AA
13	0.000	2.783	2.755	0.511	0.055	AA
14	0.000	2.759	2.747	0.523	0.023	AA
15	0.000	2.731	2.701	0.563	0.052	AA
16	0.010	2.753	2.743	0.548	0.018	AA
17	0.000	2.712	2.723	0.578	-0.019	AA
18	0.000	2.850	2.829	0.453	0.048	AA
19	0.248	2.824	2.819	0.482	0.008	AA
20	0.444	2.750	2.746	0.563	0.008	AA
21	0.060	2.740	2.732	0.568	0.014	AA
22	0.000	2.766	2.783	0.528	-0.032	AA
23	0.000	2.804	2.791	0.505	0.026	AA
24	0.000	2.790	2.775	0.522	0.029	AA
25	0.004	2.781	2.773	0.538	0.016	AA
26	0.000	2.606	2.671	0.616	-0.106	AA
27	0.000	2.627	2.681	0.611	-0.089	AA
28	0.000	2.798	2.830	0.487	-0.066	AA
29	0.000	2.766	2.806	0.515	-0.078	AA
30	0.000	2.829	2.855	0.450	-0.057	AA
31	0.000	2.842	2.862	0.447	-0.046	AA
32	0.000	2.831	2.856	0.448	-0.056	AA
33	0.000	2.802	2.830	0.475	-0.058	AA
34	0.000	2.709	2.758	0.541	-0.091	AA
35	0.000	2.615	2.682	0.585	-0.113	AA
36	0.000	2.716	2.753	0.546	-0.069	AA
37	0.000	2.645	2.678	0.593	-0.054	AA

TABLE B.2**DIF Data: Year-End Performance Indicators for Kindergarten and Ethnicity**

Indicator	<i>P</i>-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
1	0.000	2.609	2.567	0.519	0.082	AA
2	0.000	2.541	2.492	0.570	0.087	AA
3	0.000	2.584	2.622	0.526	-0.073	AA
4	0.000	2.741	2.763	0.429	-0.051	AA
5	0.000	2.683	2.676	0.484	0.015	AA
6	0.000	2.649	2.633	0.499	0.032	AA
7	0.000	2.474	2.441	0.605	0.056	AA
8	0.363	2.460	2.453	0.600	0.010	AA
9	0.026	2.725	2.737	0.444	-0.028	AA
10	0.991	2.791	2.792	0.390	-0.004	AA
11	0.000	2.717	2.686	0.462	0.066	AA
12	0.000	2.676	2.713	0.473	-0.080	AA
13	0.000	2.613	2.642	0.509	-0.057	AA
14	0.039	2.595	2.589	0.520	0.012	AA
15	0.000	2.550	2.524	0.562	0.046	AA
16	0.013	2.596	2.599	0.540	-0.005	AA
17	0.000	2.533	2.565	0.569	-0.057	AA
18	0.000	2.727	2.704	0.447	0.051	AA
19	0.125	2.692	2.685	0.474	0.015	AA
20	0.017	2.591	2.580	0.556	0.019	AA
21	0.000	2.585	2.560	0.561	0.046	AA
22	0.103	2.624	2.630	0.518	-0.011	AA
23	0.544	2.653	2.649	0.499	0.009	AA
24	0.000	2.643	2.620	0.517	0.044	AA
25	0.795	2.618	2.616	0.533	0.003	AA
26	0.000	2.430	2.450	0.611	-0.033	AA
27	0.000	2.451	2.469	0.604	-0.030	AA
28	0.416	2.684	2.681	0.483	0.008	AA
29	0.145	2.637	2.645	0.510	-0.016	AA
30	0.000	2.721	2.738	0.442	-0.038	AA
31	0.000	2.730	2.749	0.443	-0.044	AA
32	0.000	2.725	2.740	0.440	-0.035	AA
33	0.000	2.681	2.710	0.465	-0.062	AA
34	0.000	2.560	2.595	0.535	-0.067	AA
35	0.000	2.449	2.487	0.581	-0.066	AA
36	0.244	2.575	2.583	0.540	-0.014	AA
37	0.000	2.463	2.503	0.588	-0.068	AA

Table B.3**DIF Data: Year-End Performance Indicators for First Grade and Gender**

Indicator	<i>P</i>-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
1	0.405	2.700	2.703	0.542	-0.007	AA
2	0.000	2.677	2.636	0.599	0.067	AA
3	0.000	2.769	2.663	0.540	0.195	BB+
4	0.000	2.871	2.800	0.452	0.156	AA
5	0.000	2.817	2.750	0.501	0.133	AA
6	0.000	2.771	2.747	0.528	0.045	AA
7	0.000	2.705	2.623	0.612	0.133	AA
8	0.000	2.650	2.628	0.616	0.036	AA
9	0.000	2.804	2.784	0.485	0.041	AA
10	0.000	2.838	2.807	0.454	0.068	AA
11	0.000	2.807	2.792	0.485	0.031	AA
12	0.000	2.826	2.764	0.489	0.126	AA
13	0.000	2.685	2.640	0.588	0.077	AA
14	0.000	2.736	2.709	0.537	0.050	AA
15	0.000	2.684	2.644	0.599	0.067	AA
16	0.000	2.723	2.701	0.560	0.039	AA
17	0.052	2.667	2.665	0.601	0.003	AA
18	0.000	2.849	2.799	0.461	0.108	AA
19	0.000	2.839	2.827	0.465	0.025	AA
20	0.015	2.732	2.726	0.558	0.011	AA
21	0.000	2.678	2.665	0.605	0.021	AA
22	0.030	2.710	2.716	0.572	-0.011	AA
23	0.000	2.672	2.627	0.613	0.074	AA
24	0.000	2.685	2.650	0.594	0.060	AA
25	0.000	2.666	2.621	0.623	0.072	AA
26	0.000	2.710	2.785	0.530	-0.142	AA
27	0.000	2.633	2.722	0.583	-0.153	AA
28	0.000	2.788	2.842	0.471	-0.116	AA
29	0.000	2.741	2.805	0.514	-0.125	AA
30	0.000	2.575	2.653	0.613	-0.128	AA
31	0.000	2.817	2.851	0.446	-0.077	AA
32	0.000	2.827	2.860	0.436	-0.075	AA
33	0.000	2.799	2.838	0.461	-0.085	AA
34	0.000	2.664	2.728	0.561	-0.115	AA
35	0.000	2.660	2.727	0.556	-0.121	AA
36	0.000	2.698	2.753	0.539	-0.101	AA
37	0.000	2.709	2.770	0.526	-0.116	AA
38	0.000	2.702	2.749	0.550	-0.087	AA
39	0.000	2.629	2.693	0.588	-0.108	AA

Table B.4**DIF Data: Year-End Performance Indicators for First Grade and Ethnicity**

Indicator	<i>P</i>-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
1	0.000	2.560	2.510	0.540	0.093	AA
2	0.000	2.478	2.424	0.598	0.090	AA
3	0.000	2.535	2.600	0.542	-0.120	AA
4	0.000	2.700	2.734	0.454	-0.075	AA
5	0.001	2.640	2.640	0.503	0.001	AA
6	0.000	2.602	2.584	0.528	0.034	AA
7	0.000	2.491	2.429	0.612	0.101	AA
8	0.262	2.424	2.418	0.616	0.010	AA
9	0.000	2.658	2.678	0.487	-0.042	AA
10	0.006	2.701	2.715	0.454	-0.031	AA
11	0.000	2.674	2.653	0.484	0.045	AA
12	0.000	2.636	2.700	0.491	-0.131	AA
13	0.000	2.451	2.491	0.589	-0.068	AA
14	0.000	2.559	2.542	0.536	0.033	AA
15	0.000	2.469	2.452	0.597	0.027	AA
16	0.000	2.535	2.556	0.554	-0.037	AA
17	0.000	2.453	2.499	0.595	-0.078	AA
18	0.000	2.717	2.670	0.459	0.103	AA
19	0.000	2.719	2.698	0.457	0.047	AA
20	0.000	2.577	2.556	0.550	0.037	AA
21	0.000	2.496	2.471	0.598	0.042	AA
22	0.368	2.543	2.537	0.564	0.009	AA
23	0.000	2.464	2.423	0.608	0.067	AA
24	0.000	2.511	2.450	0.588	0.104	AA
25	0.000	2.455	2.414	0.618	0.065	AA
26	0.000	2.585	2.604	0.526	-0.035	AA
27	0.000	2.475	2.510	0.580	-0.061	AA
28	0.000	2.680	2.698	0.467	-0.038	AA
29	0.000	2.612	2.641	0.510	-0.057	AA
30	0.000	2.391	2.433	0.610	-0.069	AA
31	0.000	2.702	2.731	0.442	-0.067	AA
32	0.000	2.722	2.739	0.433	-0.040	AA
33	0.000	2.681	2.705	0.458	-0.051	AA
34	0.000	2.503	2.539	0.559	-0.064	AA
35	0.000	2.509	2.543	0.552	-0.062	AA
36	0.000	2.547	2.585	0.536	-0.072	AA
37	0.001	2.580	2.596	0.521	-0.031	AA
38	0.000	2.541	2.573	0.546	-0.060	AA
39	0.000	2.453	2.493	0.584	-0.069	AA

Table B.5**DIF Data: Year-End Performance Ratings for Kindergarten and Gender**

Functional Component	<i>P</i> -Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
Self-Concept	0.000	2.679	2.665	0.530	0.026	AA
Self-Control	0.000	2.858	2.783	0.463	0.163	AA
Approaches to Learning	0.000	2.715	2.694	0.569	0.037	AA
Interaction with Others	0.000	2.840	2.820	0.430	0.045	AA
Social Problem-Solving	0.000	2.783	2.755	0.511	0.055	AA
Communication	0.003	2.715	2.706	0.529	0.017	AA
Reading	0.094	2.744	2.740	0.531	0.008	AA
Writing	0.000	2.801	2.791	0.510	0.021	AA
Mathematical Processes	0.000	2.588	2.654	0.599	-0.110	AA
Number and Operations	0.000	2.755	2.798	0.499	-0.086	AA
Patterns, Relationships, and Functions	0.000	2.807	2.832	0.452	-0.056	AA
Geometry and Spatial Relations	0.000	2.778	2.808	0.467	-0.065	AA
Measurement	0.000	2.682	2.737	0.552	-0.099	AA
Data Collection and Probability	0.000	2.645	2.678	0.593	-0.054	AA

Table B.6**DIF Data: Year-End Performance Ratings for Kindergarten and Ethnicity**

Functional Component	<i>P</i> -Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
Self-Concept	0.000	2.529	2.479	0.529	0.093	AA
Self-Control	0.000	2.690	2.705	0.463	-0.032	AA
Approaches to Learning	0.000	2.519	2.503	0.568	0.028	AA
Interaction with Others	0.841	2.714	2.717	0.428	-0.006	AA
Social Problem-Solving	0.000	2.613	2.642	0.509	-0.057	AA
Communication	0.471	2.538	2.542	0.524	-0.006	AA
Reading	0.001	2.585	2.572	0.524	0.025	AA
Writing	0.001	2.651	2.638	0.505	0.027	AA
Mathematical Processes	0.000	2.416	2.436	0.594	-0.033	AA
Number and Operations	0.349	2.630	2.636	0.495	-0.012	AA
Patterns, Relationships, and Functions	0.000	2.686	2.707	0.447	-0.049	AA
Geometry and Spatial Relations	0.000	2.653	2.679	0.459	-0.057	AA
Measurement	0.000	2.528	2.555	0.547	-0.051	AA
Data Collection and Probability	0.000	2.463	2.503	0.588	-0.068	AA

Table B.7**DIF Data: Year-End Performance Ratings for First Grade and Gender**

Functional Component	<i>P</i> -Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
Self-Concept	0.000	2.646	2.624	0.545	0.041	AA
Self-Control	0.000	2.842	2.760	0.480	0.172	BB+
Approaches to Learning	0.000	2.721	2.679	0.576	0.072	AA
Interaction with Others	0.000	2.809	2.778	0.460	0.068	AA
Social Problem-Solving	0.000	2.685	2.640	0.588	0.077	AA
Communication	0.000	2.678	2.654	0.546	0.044	AA
Reading	0.000	2.707	2.693	0.552	0.025	AA
Writing	0.000	2.685	2.641	0.606	0.073	AA
Mathematical Processes	0.000	2.629	2.718	0.546	-0.164	AA
Number and Operations	0.000	2.741	2.804	0.514	-0.122	AA
Patterns, Relationships, and Functions	0.000	2.797	2.834	0.446	-0.083	AA
Geometry and Spatial Relations	0.000	2.675	2.735	0.513	-0.117	AA
Measurement	0.000	2.695	2.755	0.535	-0.111	AA
Data Collection and Probability	0.000	2.625	2.685	0.559	-0.107	AA

Table B.8**DIF Data: Year-End Performance Ratings for First Grade and Ethnicity**

Functional Component	<i>P</i> -Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
Self-Concept	0.000	2.476	2.426	0.544	0.093	AA
Self-Control	0.000	2.647	2.684	0.482	-0.076	AA
Approaches to Learning	0.000	2.514	2.485	0.576	0.051	AA
Interaction with Others	0.000	2.651	2.673	0.461	-0.047	AA
Social Problem-Solving	0.000	2.451	2.491	0.589	-0.068	AA
Communication	0.004	2.473	2.483	0.543	-0.019	AA
Reading	0.000	2.536	2.512	0.546	0.044	AA
Writing	0.000	2.484	2.435	0.600	0.081	AA
Mathematical Processes	0.000	2.483	2.516	0.543	-0.062	AA
Number and Operations	0.000	2.609	2.637	0.510	-0.055	AA
Patterns, Relationships, and Functions	0.000	2.679	2.703	0.443	-0.053	AA
Geometry and Spatial Relations	0.000	2.525	2.563	0.511	-0.074	AA
Measurement	0.000	2.548	2.580	0.531	-0.060	AA
Data Collection and Probability	0.000	2.452	2.496	0.556	-0.079	AA

WORKS CITED

- Allen, Nancy L., James E. Carlson, and Christine A. Zalanak. 1999. *The NAEP 1996 Technical Report*. Washington, DC: National Center for Education Statistics.
- Clauser, Brian E., and Kathleen M. Mazor. 1998. "An NCME Instructional Module on Using Statistical Procedures to Identify Differentially Functioning Test Items." *Educational Measurement: Issues and Practice* 17, no. 1 (spring): 31–44.
- Dorans, Neil J., Alicia P. Schmitt, and Carole A. Bleistein. 1992. "The Standardization Approach to Assessing Comprehensive Differential Item Functioning." *Journal of Educational Measurement* 29:309–19.
- HEM. 1999. *A Proposal to South Carolina Department of Education for the Readiness Assessment for First and Second Graders*. San Antonio, TX: Harcourt Educational Measurement.
- . 2002. *South Carolina Readiness Assessment Developmental Checklists District and School Test Administrator Manual 2002–2003*. San Antonio, TX: Harcourt Educational Measurement.
- Hills, J. R. 1989. "Screening for Potentially Biased Items in Testing Programs." *Educational Measurements Issues and Practice* 8, no. 4:5–11.
- Huynh, Huynh, Susan V. Prior, and Dorinda J. Gallant-Taylor. 2004. *Technical Documentation for the 2002 South Carolina Readiness Assessment of Kindergarten and Grade-One Students*. Columbia: South Carolina Department of Education.
- Monrad, Diane M., and Garrett K. Mandeville. 1996. *Evaluation of the Work Sampling System Field Test*. Columbia: South Carolina Educational Policy Center.
- Porchea, Sameano F., and James F. Casteel. 2003. "The South Carolina Readiness Assessment Profile Interactive (SCRAPI)." Paper presented at annual meeting of the National Council on Measurement in Education. Chicago, IL, April 22–24.
- Rebus and SDE. 2002. *South Carolina Readiness Assessment Kindergarten and First Grade Developmental Guidelines*. New York: Pearson Early Learning.
- Stokes, Maura E., Charles S. Davis, and Gary G. Koch. 2001. *Categorical Data Analysis Using the SAS System*. 2nd ed. New York: Wiley.
- Zwick, Rebecca, and Kadriye Ercikan. 1989. "Analysis of Differential Item Functioning in the NAEP History Assessment." *Journal of Educational Measurement* 26:55–66.

- Zwick, Rebecca, and Dorothy T. Thayer. 1996. "Evaluating the Magnitude of Differential Item Functioning in Polytomous Items." *Journal of Educational and Behavioral Statistics* 21:187–201.
- Zwick, Rebecca, Dorothy T. Thayer, and John Mazzeo. 1997. "Descriptive and Inferential Procedures for Assessing DIF in Polytomous Items." *Applied Measurement in Education* 10:321–44.